DOCUMENT RESUME

ED 466 353 SE 065 729

AUTHOR	Kilpatrick, Jeremy
TITLE	State Proficiency Testing in Mathematics. ERIC Digest.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, OH.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-SE-01-07
PUB DATE	2001-12-00
NOTE	4p.
CONTRACT	ED-99-CO-0024
AVAILABLE FROM	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, 1929 Kenny Road, Columbus, OH 43210-1080. Tel: 800-276-0462 (Toll Free); Fax: 614-292-0263; e-mail: ericse@eric.org. For full text: http://www.ericse.org.
PUB TYPE	ERIC Publications (071) -- ERIC Digests in Full Text (073)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Standards; *Achievement Tests; Elementary Secondary Education; Evaluation Methods; High Stakes Tests; Mathematics Education; *Standardized Tests; *State Standards
IDENTIFIERS	ERIC Digests

ABSTRACT

This digest describes state policies and teacher dilemmas in using proficiency tests and standardized assessment. Currently, 19 states require mandated assessment for high school graduation and several other states are considering such requirements. Standards-based assessment needs to be accompanied by a clear set of grade-level goals so teachers, parents, and others can work together to help all children in a school achieve those goals. (Contains 12 references.) (YDS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 466 353

ERIC Digest

State Proficiency Testing in Mathematics

by Jeremy Kilpatrick

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SE065729

BEST COPY AVAILABLE

ERIC Educational Resources Information Center

State Proficiency Testing in Mathematics

Jeremy Kilpatrick

December 2001

DIGEST

EDO-SE-01-07

Clearinghouse for Science, Mathematics, and Environmental Education

In recent years many states have developed their own assessments of student learning in mathematics, usually aligned with state standards or curriculum frameworks. Many of these assessments are intended to have *high stakes*: financial or other consequences for districts, schools, teachers, or individual students. In some cases, teacher promotions or high school diplomas may depend on students achieving passing scores. As of 1998, 48 states and the District of Columbia had instituted testing programs, typically at grades 4, 8, and 11 (Council of Chief State School Officers, 1998).

Many states report the results of high-stakes assessments by school or by district to identify places most in need of improvement. State responses to assessment results vary. Some states have the authority to close, take over, or "reconstitute" a failing school, but to date only a few states have ever used such sanctions (Jerald, Curran, & Boser, 1999). Florida awards additional funds to schools that perform near the bottom or near the top of the range (Sandham, 1999). When schools or districts with poor results do not show sufficiently rapid improvement, some states revoke accreditations, close the schools, seize control of the schools, or grant vouchers that enable students to enroll elsewhere.

Currently, 19 states require students to pass a mandated assessment in order to graduate from high school, and several other states are phasing in such a requirement (Gehring, 2000). In response to calls for an end to social promotion, some states and districts have begun requiring grade-level mastery tests for promotion, typically in grades 4 and 8. Interestingly, there is some evidence suggesting an inverse relationship between statewide testing policies and student achievement in mathematics:

> Among the 12 highest-scoring states in 8th grade mathematics in 1996, ... none had mandatory statewide testing programs in place during the 1980s or early 1990s. Only two of the top 12 states in the 4th grade mathematics had statewide programs prior to 1995. By contrast, among the 12 lowest-scoring states, ... 10 had extensive student testing programs in place prior to 1990, some of which were associated with highly specified state curricula and an extensive menu of rewards and sanctions (Darling-Hammond, 1999, p. 33).

Responses to Trial Runs

To give teachers, students, parents, and others sufficient time to prepare for high-stakes assessments, states typically administer them for several years before the consequences take effect. During the trial period, failure rates are sometimes alarmingly high. In Arizona, for example, only 1 in 10 sophomores passed the mathematics test first given in the spring of 1999. During the same period, only 7% of Virginia schools were able to achieve the 70% passing rate, which was to become a condition for accreditation in 2007. In response to these initial results, some states have begun relaxing their expectations, reconsidering the tests, or withdrawing them altogether. Wisconsin, for example, yielded to pressure from parents and withdrew its high school graduation test. Massachusetts and New York set lower passing scores for their exams (Steinberg, 1999).

Most states report the level of student performance on their assessments by setting so-called "cut scores" to define categories with such labels as *advanced, proficient, needs improvement,* and *failing* (Elmore & Rothman, 1999), terms similar to those used in the National Assessment of Educational Progress (NAEP): *advanced, proficient,* and *basic*. When results on state assessments are compared with the state NAEP results, the proportions of students reaching the proficient level are often higher (Archer, 1997). Some have concluded from this discrepancy that most state tests do not reflect sufficiently high expectations (Musick, 1997). Others argue that minimum competence and high expectations are different goals that cannot be measured by the same assessment and certainly not with the same cut scores. Thus, the results appear discrepant because the same categories are used to describe performance on assessments with very different goals.

Standardized Assessments

Many states and school districts also administer standardized tests which may or may not coincide with state assessments. Commercially published standardized achievement tests are quite variable in terms of the topics covered and the degree that topics are emphasized at each grade level (Romberg & Wilson, 1992, and they are frequently not aligned with the teaching materials used in districts or with district goals. This misalignment further dilutes teaching efforts, as teachers must add to their long lists of goals and topics to be covered.

Most standardized tests might be called "comparison" tests because their function is to rank order students, schools, and districts. Items are chosen to range widely in difficulty so as to disperse students' scores, allowing half the students to be classified as "below average" and the other half as "above average." The tests do not include many items that only a few students get right or wrong because such items do not help distinguish among students. The omission of such items may lead to some important aspects of mathematics not being tested, but for tests designed to make comparisons, the omissions are necessary.

In contrast, if the purpose of a test is to assess whether students have met specific goals, test designers can choose items to span the important mathematics to be learned, and cut scores can be set to indicate various levels of proficiency. Students and teachers know where to focus their efforts and prepare for tests with the goals in mind. If students have learned well, large proportions of them can achieve high proficiency; there is no need to label half of them as below average or to rank them in any way.

There has traditionally been a level of secrecy about standardized tests so questions can be reused. In recent years this practice has come under fire. If students are to

reach publicly accepted standards, the argument goes, they need to know what types of performance will be expected of them (Rothman, 1995, p. 5). Legally and ethically, when the stakes are high, students should be provided with sample assessments or at least sample items that are representative of the actual assessments (Heubert & Hauser, 1998).

A Dilemma for Teachers

The movementto hold schools accountable for student performance has resulted in increased high-stakes testing of "minimum competencies" in mathematics. Many states give competency tests at several grade levels, including high school exit exams, and performance on such tests has often been considerably below what was anticipated or desired. Meanwhile, many districts continue to use standardized tests that are not necessarily aligned with textbooks, state goals, or state competency tests. This combination of standardized comparison tests and state competency tests can overwhelm teachers who have to prepare students for two kinds of tests about which they often know very little.

State competency tests are often given first at a grade level at which many students are already far behind in mathematics and likely to have difficulty in catching up. If such tests are to be used, they need to be accompanied in earlier grades—and throughout all grades—by other assessments that would enable teachers to make instruction more effective. In particular, such assessments could identify students who are not achieving and need special help so that they do not fall further behind. This linking of assessment to instructional efforts is consistent with a recent NRC report (Elmore & Rothman, 1999) which includes the following two central recommendations:

- Teachers should administer assessments frequently and regularly in classrooms for the purpose of monitoring individual students' performance and adapting instruction to improve their performance. (p. 47)
- Teachers should monitor the progress of individual children in grades pre-K to 3 to improve the quality and appropriateness of instruction. Such assessments should be conducted at multiple points in time, in children's natural settings, and should use direct assessments, portfolios, checklists, and other work sampling devices. (p. 53)

Conclusion

The current national focus on standards-based testing is an improvement over the past focus on comparison testing. But standards-based assessment needs to be accompanied by a clear set of grade-level goals so teachers, parents, and others can work together to help all children in a school achieve those goals. Continuing informal assessments throughout the year can help teachers adjust their teaching and identify students who need additional help. More such help might be available if money formerly spent on comparison testing were reallocated to help children learn.

References

Archer, J. (1997, January 15). States struggle to ensure data make the grade. *Education Week* [On-line]. Available: http://www.edweek.com/ew/1997/16data.h16

Council of Chief State School Officers, State Education Assessment Center. (1998). *Key state education policies on K-12 education.* Washington, DC: Author. Available: http://publications.ccsso.org/ccsso/publication_detail.cfm?PID=187

Darling-Hammond, L. (1999, December). *Teacher quality and student achievement: A review of state policy evidence.* Seattle, WA: Center for the Study of Teaching and Policy. Available: http://depts.washington.edu/ctpmail/ or http://olam.ed.asu.edu/epaa/v8n1/

Elmore, R. F., & Rothman, R. (Eds.). (1999). *Teaching, testing, and learning: A guide for states and school districts.* Washington, DC: National Academy Press. Available: http://books.nap.edu/catalog/9609.html

Gehring, J. (2000, February 2). "High stakes" exams seen as test for voc. ed. *Education Week* [On-line]. Available: http://www.edweek.org/ew/ewstory.cfm?slug=21voctest.h19

Heubert, J. P., & Hauser, R. M. (Eds.). (1998). *High stakes: Testing for tracking, promotion, and graduation.* Washington, DC: National Academy Press. Available: http://books.nap.edu/catalog/6336.html

Jerald, C. D., Curran, B. K., & Boser, U. (1999, January 11). The state of the states [Quality Counts '99]. *Education Week*, pp. 106-123. Available: http://www.edweek.org/sreports/qc99/states/indicators/in-intro.htm

Musick, M. (1997). *Setting education standards high enough.* Atlanta, GA: Southern Regional Education Board. Available: http://www.sreb.org/main/highschools/accountability/settingstandardshigh.asp

Romberg, T. A., & Wilson, L. D. (1992). Alignment of tests with the standards. *Arithmetic Teacher, 40*(1), 18-22.

Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform.* San Francisco: Jossey-Bass.

Sandham, J. L. (1999, July 14). In first for states, Florida releases graded "report cards" for schools. *Education Week* [On-line]. Available: http://edweek.org/ew/1999/42fla.h18

Steinberg, J. (1999, December 3). Academic standards eased as a fear of failure spreads. *The New York Times*, p. A1.

Additional Web Resources

Adding It Up: Helping Children Learn Mathematics by Jeremy Kilpatrick, Jane Swafford, & Bradford Findell (Editors), Online publication of the National Academy Press. http://www.nap.edu/catalog/9822.html

ERIC Resources

Additional materials pertaining to high stakes testing in mathematics are described in the ERIC database. Search the ERIC database at http://ericir.syr.edu/Eric/adv_search.shtml, and use ERIC Descriptors such as "mathematics tests" and "high stakes tests."

SE 065 729 This digest is in the public domain and may be freely reproduced. EDO-SE-01-07

This digest was funded by the Office of Educational Research and Improvement, U.S. Department of Education, under contract no. ED-99-CO-0024. Opinions expressed in this digest do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education.

The Educational Resources Information Center is a nationwide information system initiated in 1966 by the U.S. Department of Education. ERIC has developed the largest and most frequently used education-related database in the world. For information, call 1-800-538-3742.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").